LA MACCHINA DELLA MENTE

Parte I[a]. Il cervello e il calcolatore.

Autore: Renato Nobili
Dipartimento di Fisica "G.Galilei" – Università di Padova.

Sommario: In questa prima parte del saggio sul funzionamento del cervello tratterò alcuni argomenti riguardanti l'intelligenza artificiale. Dopo una breve analisi storica del paradigma logico booleano introdurrò alcune questioni teoriche che sono essenziali per comprendere che cosa sia la capacità autoriflessiva del pensiero e cercherò di spiegare le ragioni che portano a considerare il cervello umano come una *macchina ricorsiva universale*. Metterò infine a confronto i principi di funzionamento del calcolo seriale con quelli del calcolo parallelo per cercare di comprendere se vi siano sostanziali differenze tra le possibilità di elaborazione dei calcolatori e quelle dei cervelli animali.

In una seconda parte descriverò brevemente i principali modelli e paradigmi che hanno ispirato le neuroscienze negli anni '80 e '90, principalmente allo scopo di evidenziare i limiti e gli errori che hanno causato il loro fallimento.

In una terza parte cercherò di fornire un'idea dei principali aspetti anatomici e funzionali del cervello umano, allo scopo di illustrare le idee più interessanti che sono state avanzate nel corso degli anni recenti e le prospettive che si sono aperte per la comprensione del suo funzionamento. Data la complessità, la vastità dell'argomento e la limitatezza delle mie conoscenze in campo neuroscientifico – dovute soprattutto al fatto che non sono uno specialista della materia - il contenuto della terza parte sarà inevitabilmente assai incompleto e limitato a pochi argomenti scelti secondo i miei gusti personali o seguendo i miei particolari percorsi di esplorazione.

INDICE DELLA PRIMA PARTE

1.1. Il sogno di Boole

Per quasi tutto il quarantennio successivo all'apparizione dell'articolo *A Logical Calculus of Ideas Immanent in Nervous Activity* di McCulloch e Pitts (1943) il neurone è stato modellizzato come un *decisore a soglia*, ossia un dispositivo che genera in uscita un segnale a gradino quando una somma algebrica di segnali positivi o negativi applicati in ingresso supera un certo valore chiamato appunto soglia (Fig.1). Per tutto quel periodo la complessità dei processi nervosi fu sottovalutata e le possibilità del calcolo elettronico sopravvalutate.

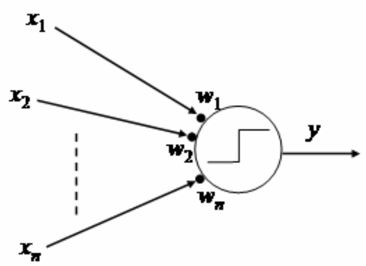

Fig.1. Il neurone come dispositivo a soglia:
Se $x_1 + x_2 + ... + x_N >$ *soglia,* $y=1$, *altrimenti* $y = 0$.
w_1, w_2, ..., w_n = coefficieni di trasmissione (pesi) delle sinapsi.

Per tali ragioni il funzionamento del neurone poteva apparire tanto elementare quanto quello di una valvola termoionica o di un transistor, la cui scoperta, è il caso di ricordare, risale agli inizi degli anni 50. Poiché allora si ignorava che il tessuto di sostegno e di alimentazione della rete nervosa (la glia) scambia segnali con i neuroni, svolgendo per questi importantissime funzioni di regolazione, lo scambio di segnali poteva sembrare unicamente effettuato dalle cellule nervose. Era pertanto naturale pensare che i processi mentali fossero riconducibili all'interazione di unità funzionali simili ai componenti elementari di un circuito elettronico (Arbib, 1965). L'analogia sembrava tanto più corretta in quanto si riteneva che i segnali nervosi si propagassero elettricamente attraverso i contatti sinaptici intesi come punti di saldatura tra le cellule nervose. La localizzazione della memoria appariva invece problematica. Alcuni ritenevano che l'informazione nervosa fosse memorizzata nei valori di soglia dei neuroni, altri che fosse codificata in molecole contenute nei neuroni, magari nello stesso DNA. Delle diverse ipotesi allora formulate rimane valida ancora oggi quella di Hebb (1949), che fu il primo ad ipotizzare che la memoria sia dovuta alla formazione o all'alterazione dei contatti sinaptici che mettono in comunicazione le cellule di una popolazione neuronale. Per capire la differenza tra la visione di allora e quella dei nostri giorni basta pensare che oggi la complessità funzionale di un singolo neurone è paragonata da alcuni a quella di un microprocessore. In realtà il paragone è in parte esagerato in parte sbagliato perché i singoli neuroni non possiedono notevoli capacità di calcolo e inoltre sono variamente specializzati per lo svolgimento di funzioni diverse.

L'aspetto più suggestivo di quella visione tanto ottimistica era fornito dal modello della risposta neuronale di tipo *tutto o niente*. Questo semplice principio di funzionamento trovava la sua legittimazione teorica nella rappresentazione in cifre binarie o bit (*binary digit*), ossia mediante successioni di 0 e 1, dell'informazione intesa come una sequenza di decisioni dicotomiche (Shannon, 1949). Tutto ciò corrispondeva pienamente a quella che poteva sembrare la profetica eredità culturale di George Boole (1853) fondatore del calcolo logico: l'idea che le regole di formazione del linguaggio logico [*congiunzione* (AND), *disgiunzione* (OR), *negazione* (NOT) e loro combinazioni], in definitiva le stesse operazioni fondamentali del pensiero, costituissero una sorta di algebra operante su variabili proposizionali a due valori: *tutto*=1= *vero*; *nulla* = 0 = *falso*.

Del resto, se si assume che i costituenti elementari delle reti nervose funzionino come decisori a soglia, appare naturale ipotizzare che nel cervello avvengano processi analoghi a quelli del calcolo di Boole.

Si può immaginare quale sia stato il divertimento di McCulloch e Pitts, e dopo di loro di Kleene (1956) che riprese la questione in modo più completo e rigoroso, nel riuscire a dimostrare che tutte le operazioni del calcolo booleano possono essere eseguite combinando decisori a soglia temporalmente ritardati (Fig.2).

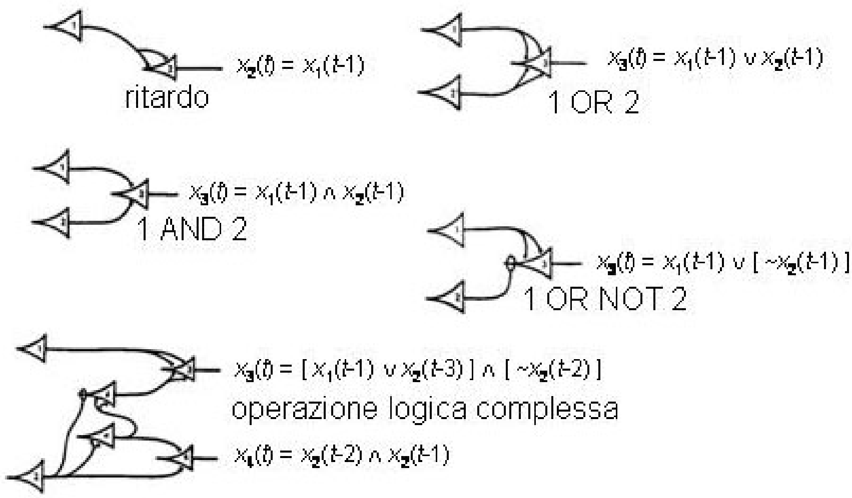

Figura 2. Alcuni circuiti neuronali descrtti da McCulloch e Pitts, i loro equivalenti operazionali boleani e le corrispondenti formule logiche. Da McCulloch & Pitts (1943).

D'altronde, poiché il funzionamento di un decisore a soglia può essere approssimato tanto bene quanto si vuole da un circuito formato da operatori booleani elementari e dispositivi di ritardo temporale, si deduceva che ogni elaborazione d'informazione effettuabile mediante decisori a soglia poteva essere simulata da una rete sufficientemente complessa di operatori booleani. Insomma, si deduceva che se le reti neurali di McCulloch e Pitts fossero davvero buoni modelli di reti nervose reali, i cervelli animali e i calcolatori digitali sarebbero due modi perfettamente equivalenti di processare l'informazione, con tutte le conseguenze che si possono immaginare, anche sul piano filosofico. L'impetuoso sviluppo delle tecnologie informatiche degli anni '60 contribuì ad alimentare ulteriormente quest'illusione di semplicità (De Luca e Ricciardi, 1981).

Fig.3. Kurt Gödel (1906-1978)

Fig.4. Alan Mathison Turing (1912-1954).

In realtà, come dimostrò Alan Turing nel 1936, per raggiungere grandi capacità di calcolo una rete di operatori logici da sola non basta, ma servono anche dispositivi capaci di eseguire operazioni di tipo aritmetico e di scrittura, lettura e cancellazione di dati, in modo da permettere l'esecuzione di procedure iterative (o per dire meglio *ricorsive*) istruite da programmi sostituibili. Mediante procedure ricorsive applicate a insiemi di dati è possibile implementare calcoli matematici d'ogni genere e complessità e raggiungere così l'*universalità computazionale*. Vale a dire la capacità di simulare qualsiasi altro calcolatore.

La questione delle possibilità di calcolo dei processi algoritmici ricorsivi si collegò così in modo naturale alla stupenda problematica relativa al rapporto tra la *logica* e l'*aritmetica*: in particolare ai teoremi di decidibilità e indecidibilità di Kurt Gödel (1930, 1931) e alla teoria del rapporto tra linguaggio e metalinguaggio di Alfred Tarski (1936), ecc. Sembrò allora che si potesse effettuare anche la "quadratura" del cervello: che tutto ciò che accade nel cervello possa essere simulato da un calcolatore sufficientemente complesso, magari con l'aggiunta di una sorgente fisica di dati casuali dato che questi non possono essere generati da processi di calcolo deterministici.

Fig. 5. John von Neumann (1901-1956) e il primo calcolatore

Semmai il problema era di spiegare come il cervello potesse eseguire calcoli precisi con dispositivi imprecisi e deperibili come sono i neuroni. A questo proposito von Neumann dimostrò che questo problema poteva essere risolto aumentando la ridondanza dei circuiti neuronali.

Sulla base delle idee di Turing, von Neumann teorizzò e costruì l'ENIAC, il primo calcolatore elettronico universale. L'indagine sui fondamenti del pensiero sembrò chiudersi così in un circolo perfetto: il pensiero è logico perché gli elementi funzionali del cervello sono simili a quelli dei calcolatori. Si poteva pretendere di più? Quali dubbi avrebbero potuto sorgere circa la validità di questo approccio dato che molti altri matematici come Shannon, Moore, Arbib, ecc. contribuirono a legittimarlo elevandolo al rango di una teoria sistematica e completa? Tale concezione ha informato fino a tempi recenti la ricerca sull'*intelligenza artificiale* (IA), a proposito della quale merita ricordare i nomi di Minsky, Newell, Simon.

Purtroppo, i risultati in questo campo sono stati tanto deludenti quanto le speranze erano state promettenti. In 40 anni di intense ricerche, il software prodotto dagli studiosi di IA non ha mai fornito qualcosa che assomigli vagamente a un'implementazione del pensiero umano. A cosa furono dovute queste difficoltà? In questo saggio cercherò di rispondere a questa domanda.

1.2. Gödel, Turing e von Neumann

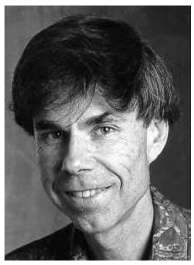

Fig. 6. Douglas Hofstadter (1945).

Alcuni logici e teorici dell'IA hanno nutrito per alcuni decenni un ambizioso progetto: spiegare come un processo di elaborazione dell'informazione possa essere capace di "autoriflessione", di riprodurre, cioè, la capacità della mente di pensare il proprio pensato, in altri termini di generare l'*autocoscienza*.

Questo argomento è stato divulgato in forma avvincente e suggestiva da Douglas Hofstadter nel libro *Gödel, Escher, Bach* (1979). L'autore trasse ispirazione dalle proprietà autoreferenziali dell'aritmetica scoperte da Gödel; ma anche, sebbene l'autore non lo dichiari, da quella, del tutto affine, della teoria degli macchine dotate di capacità costruttive (automi) e persino capaci di autoriprodursi, delineata da John von Neumann in una serie di conferenze e appunti (1966) raccolti dai suoi collaboratori tra il 1949 e il 1956; anno, questo, della morte del grande matematico.

Nella visioni di Gödel e di von Neumann le possibilità, rispettivamente, dell'autoreferenza e dell'autoriproduzione, dipendono in modo critico dalla condizione che il sistema considerato, precisamente l'algoritmo aritmetico o l'automa costruttore, sia capace di fare tutto quello che può fare un altro sistema simile e che pertanto sia capace di esibire comportamenti o produrre risultati di tutti i tipi e di qualsiasi grado di complessità, che sia, cioè, *universale*. Tuttavia, affinché ciò si verifichi, il sistema deve poter svolgere un certo numero di funzioni essenziali, ciò che comporta un certo grado di complessità strutturale. Se una sola di queste funzioni venisse a mancare il sistema perderebbe l'universalità e non sarebbe più in grado di autoreferenziarsi o di autoriprodursi. Ciò si spiega considerando che le procedure autoreferenziali e i processi di autoriproduzione non sono altro, in ultima analisi, che particolari processi computazionali o costruttivi.

Questo argomento è così importante da meritare una digressione piuttosto ampia. L'acquisizione dei concetti che ora cercherò di presentare, purtroppo in forma approssimativa e incompleta, richiedono un certo sforzo di attenzione da parte del lettore, ma la fatica ne vale la pena perché si tratta di un nuovo campo del sapere, per lo più ancora ignorato dalla cultura generale, che si sta propagando in tutte le direzioni nel territorio delle scienze della complessità e che oltrepassa il livello della cultura scientifica del secolo scorso, che era principalmente basata su idee e concetti mutuati dalla fisica. Per descrivere la sequenza logica dei passaggi necessari per acquisire questi concetti userò gli ordinali latini seguiti da titoli indicativi degli argomenti e scriverò in corsivo le definizioni più importanti.

I. Algoritmi e processi.

L'aritmetica e l'algebra elementare sono due *algoritmi* molto popolari. Anche le regole per la costruzione di figure geometriche mediante riga e compasso costituiscono un algoritmo, sebbene in questo caso si tratti di operazioni grafiche piuttosto che aritmetiche e algebriche.

Con la massima generalità possiamo dire che un algoritmo è un insieme di regole procedurali che si applicano ai dati di un certo insieme per produrre risultati che vanno ricollocati nello stesso insieme. Una sequenza di operazioni applicate a uno stesso insieme di dati costituisce un processo. Un processo si dice *ricorsivo* se la sequenza procede in modo deterministico attraverso cicli di calcolo istruiti da un programma, in modo che ad ogni ciclo possano essere utilizzati sia i dati iniziali sia quelli prodotti nel cicli precedenti.

Un processo ricorsivo può arrestarsi ad un certo istante, quando il risultato desiderato è stato ottenuto, oppure continuare idealmente per sempre al fine di produrre un risultato sempre meglio approssimato. Nella pratica matematica gli algoritmi sono elaborazioni simboliche che permettono di produrre entità matematiche, geometriche e fisiche utili alla scienza.

***II. Universalità computazionale*.**

Diciamo che un algoritmo ne *interpreta* un altro se, operando su un insieme di dati e simboli con metodi propri, generalmente diversi da quelli dell'altro, è capace di produrre risultati equivalenti a quelli dell'altro. Brevemente, se è in grado di *simulare* i comportamenti dell'altro. Un algoritmo che non sia abbastanza complesso riesce, al più, ad interpretare algoritmi simili o più semplici. Ad esempio, mediante l'uso di esponenziali e logaritmi è possibile interpretare moltiplicazioni e divisioni in termini di addizioni e sottrazioni. Il vecchio regolo dei geometri, ora caduto in disuso, serviva proprio a questo.

Un algoritmo capace di interpretare qualsiasi altro algoritmo si dice *universale*. Allo stesso modo, si dice *universale* anche un calcolatore capace di simulare qualunque altro calcolatore. La scoperta che dalla combinazione di operazioni logiche e aritmetiche si può generare un algoritmo universale è implicita nei lavori di Gödel ed è stata resa esplicita da Turing con la dimostrazione dell'esistenza di calcolatori universali. Turing ha descritto come si possa costruire una macchina capace di simulare le procedure di calcolo di qualunque altra macchina calcolatrice, anche di complessità assai maggiore.

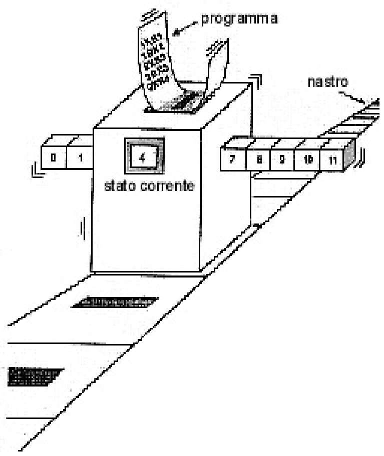

Fig.7. Macchina di Turing.

La macchina di Turing è dotata di una semplice testina mobile capace di spostarsi su un nastro di lunghezza illimitata e di leggere e scrivere su questo nastro un certo numero di simboli, memorizzando temporaneamente i dati appena letti, la posizione corrente della testina e lo stato interno della macchina in alcuni registri di piccola memoria (Fig.7). La testina si sposta avanti e indietro seguendo istruzioni scritte su una parte del nastro, che funziona come *programma*, e scrive i risultati su un'altra parte del nastro che funziona pertanto come *memoria di lavoro*. Le istruzioni sono del tipo "*se hai letto questo dato e lo stato interno è questo allora fai questo, altrimenti fai quest'altro*".
Lo schema generale delle funzioni svolte da una macchina di Turing è illustrato in Fig.8, esso corrisponde abbastanza bene anche allo schema di funzionamento di un comune calcolatore da tavolo.

L'aspetto rilevante della teoria di Turing è che l'*universalità computazionale equivale alla capacità di eseguire ogni possibile sorta di calcolo logico e/o aritmetico che si possa immaginare*, il prezzo da pagare essendo solo il tempo di calcolo. Si può perciò comprendere quanta importanza abbia avuto per la matematica la scoperta che *l'aritmetica è un algoritmo universale*. È questa la ragione per cui tutte le procedure matematiche, per esempio anche quelle puramente geometriche e persino quelle della fisica teorica, purché opportunamente interpretate, sono riconducibili, in ultima analisi, a puri e semplici calcoli logico-aritmetici. Ad esempio la rappresentazione aritmetica del teorema di Pitagora è $x^2+y^2 = z^2$.

Il teorema più importante di Turing è quello detto dell'*halting*: esso afferma che, in generale, nessuna macchina universale può prevedere se un'altra macchina universale terminerà il suo calcolo se non simulando più velocemente lo stesso calcolo.

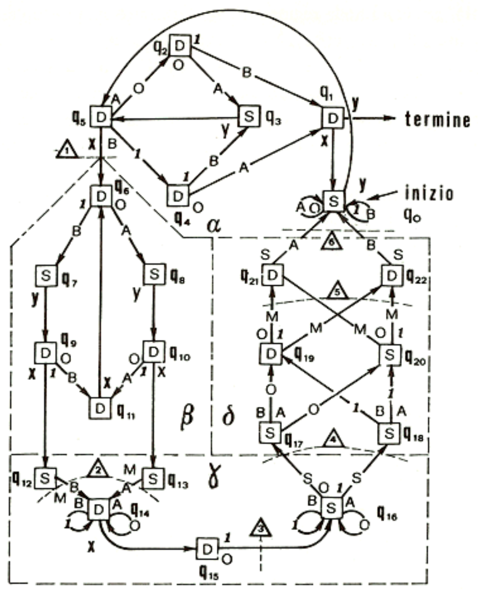

Fig.8. Schema funzionale di una macchina di Turing universale. Da Aiello et al. (1976).

III. La logica.

Ogni proposizione di un linguaggio logico, sia esso matematico o no, possiede un valore di verità: nella logica booleana la proposizione può essere *vera* o *falsa*. Combinando proposizioni logiche mediante operazioni booleane (AND, OR, NOT) si ottengono nuove proposizioni di cui si può stabilire la verità o la falsità sulla base delle tabelle di verità o falsità delle proposizioni componenti. Tutte le proposizioni che si possono ottenere mediante operazioni booleane applicate ad un insieme di proposizioni di base, indipendentemente dalla verità o falsità di queste, si dicono *proposizioni ben formate* (nel caso della geometria si potrebbe dire *figure ben disegnate*).

Ogni teoria logica parte da un certo numero di proposizioni di base, dette *assiomi*, e si propone di dimostrare altre proposizioni, dette *teoremi*. Gli assiomi sono proposizioni ben formate assunte *vere* per definizione. Lo scopo di una dimostrazione è di dedurre la verità o la falsità di una proposizione ben formata a partire della verità degli assiomi. I teoremi interessanti sono quelli che affermano verità particolarmente importanti per gli usi che se ne possono fare. Nel corso di una dimostrazione, certe proposizioni complesse possono essere rimpiazzate da altre più semplici o mediante l'introduzione di opportune definizioni o l'utilizzo di operazioni e regole logiche. In pratica, le operazioni e le regole logiche formalizzano il ragionamento matematico. Esse governano una successione di passaggi (implicazioni logiche) alcuni dei quali hanno l'effetto di aumentare la lunghezza della proposizione corrente (nel caso della geometria la complessità della figura corrente), altri di diminuirla, fino a produrre, come risultato finale, una proposizione (una figura) che è giudicata interessante. Le diminuzioni di lunghezza possono ottenersi applicando le regole d'inferenza del sistema assiomatico - le quali in pratica stabiliscono che certe formule possono essere rimpiazzate da altre meno complesse - oppure applicando le regole fondamentali della stessa logica, ad esempio il *modus ponens*, che permette di rimpiazzare la formula *"se A allora B"* semplicemente con "*B*" (nel caso della geometria è lecito cancellare le parti del disegno che sono già state usate per produrre un certo risultato). Poiché il giudizio sul valore di una proposizione è esterno al contesto delle proposizioni della teoria, si comprende che in realtà ogni teoria logica è subordinata ad un sistema interpretativo più ampio di quello formato dalla sola teoria.

Esiste tuttavia una grande differenza tra la logica di un linguaggio che descrive un mondo finito di oggetti e quella di un linguaggio matematico capace di descrivere, ad esempio, le possibilità dell'aritmetica. In questo secondo caso, l'universo del discorso matematico contiene un numero infinito di entità, che spesso consistono di variabili o incognite di cui si vuole determinare il valore o le proprietà. Ciò succede, ad esempio, quando si vuole trovare la soluzione di un'equazione. In queste circostanze il linguaggio logico ha bisogno di ipotizzare o enunciare l'*esistenza* di un oggetto entro un insieme infinito di possibili oggetti o di riferirsi a *tutti gli oggetti* di un insieme infinito. A questo proposito è importante considerare che, sebbene i dati della conoscenza e gli atti concreti dell'agire umano siano finiti, tuttavia essi si inscrivono in domini del *possibile* infiniti: infinite possibilità fisiche, infinite possibilità del pensiero, infinite posizioni che possiamo occupare nello spazio ecc.

Gli enunciati di esistenza e i riferimenti alle totalità si chiamano *quantificatori esistenziali*. Una logica che possa fare a meno dei quantificatori esistenziali si dice del *primo ordine*, una che non ne può fare a meno si dice del *secondo ordine*. In pratica, una logica del primo ordine si presenta come una collezione di proposizioni le cui verità o falsità possono essere analizzate una dopo l'altra in modo ordinato. La logica dell'aritmetica, avendo a che fare con insiemi infiniti e problemi di esistenza di oggetti entro insiemi infiniti, è del secondo ordine. Il teorema di indecidibilità di Gödel, che dimostra l'esistenza di proposizioni vere, ma di cui non si può dimostrare la verità, si presenta solo nella logica del secondo ordine. In definitiva, la logica del primo ordine si applica allo studio di mondi semplici e ben ordinati, di insiemi descrivibili, come le cose che possiamo direttamente osservare, mentre quella del secondo ordine si applica a processi e comportamenti che possono evolvere in modo imprevedibile, a insiemi di possibilità non descrivibili. Questo fatto si collega direttamente al teorema dell'*halting* di Turing.

IV. La doppia natura della matematica.

Da un punto di vista puramente formale, ogni algoritmo possiede una struttura che può essere codificata in forma assiomatica, in modo tale che le proprietà generali dell'algoritmo siano deducibili mediante dimostrazioni di teoremi. Dicendo questo non si esclude che esistano teoremi non dimostrabili. Per esempio, l'algoritmo geometrico, che fa materialmente uso di riga, matita e compasso, ha la struttura descritta in forma assiomatica dalla geometria euclidea intesa come sistema logico finalizzato alla dimostrazione delle verità geometriche (ad esempio il teorema di Pitagora).

La teoria assiomatica dell'aritmetica è stata formulata da Peano e Dedekind nella seconda metà dell'800. Essa si basa sulle nozioni primitive di *numero* e *successore* e sull'enunciazione, nei termini di queste, degli assiomi e delle regole d'inferenza che istituiscono la possibilità di eseguire quante si vogliano addizioni e moltiplicazioni, persino infinite operazioni e anche il diritto di enunciare una verità generale mediante una concatenazione infinita di verità particolari (*assioma d'induzione*). La formalizzazione logica dell'algoritmo aritmetico, esteso ai numeri reali, non è altro che l'algebra elementare (il principio d'induzione applicato alla geometria è implicito nel metodo di esaustione di Archimede).

Dunque un algoritmo può essere descritto in due modi diversi: 1) come un repertorio di operazioni che si possono eseguire concretamente nel corso del tempo su oggetti di un certo tipo; 2) come un sistema di assiomi che descrivono le regole di calcolo e le proprietà delle classi di oggetti su cui l'algoritmo può operare. Per fissare le idee limitiamoci all'algoritmo aritmetico. Nel primo modo esso può essere descritto da un linguaggio che ha come universo del discorso *i numeri interi particolari* e le operazioni da eseguire su questi. Possiamo chiamarlo l'*universo del linguaggio aritmetico*. Nel secondo modo esso è descritto da un linguaggio che ha come universo del discorso il linguaggio stesso dell'algoritmo aritmetico, e che pertanto possiamo definire *metalinguaggio* del linguaggio aritmetico. Gli oggetti di questo metalinguaggio sono le *proprietà generali dei numeri interi*, o, in altri termini, *le classi di numeri interi*, giacché la caratterizzazione di una proprietà equivale alla definizione di una classe, e delle operazioni possibili su di essi. In modo analogo, ogni altro algoritmo si presenta, per un verso, come un sistema di procedure atte a trasformare concretamente un insieme di dati particolari in un altro insieme di dati particolari; per l'altro come una struttura formale astratta che definisce ed elabora una collezione infinita di proprietà generali.

La distinzione dell'aspetto *operazionale* da quello *logico* caratterizza in modo dualistico l'intera matematica. Storicamente, tra i due aspetti si è stabilita una certa rivalità che si è manifestata, prima, nel tentativo, perseguito principalmente da Gottlob Frege (1879-1903), di ridurre l'aritmetica alla logica, poi nell'indagine di Gödel volta alla ricerca delle difficoltà logiche che s'incontrano nel descrivere le proprietà dell'algoritmo aritmetico. Il confronto dei due punti di vista ha avuto l'effetto di ampliare oltre i limiti delle loro formulazioni originarie prima la logica (attraverso i lavori di Whithead, Russel, Zermelo, Fraenkel, Hilbert, Tarsky ecc.) e poi la teoria degli algoritmi (con i contributi di Church, Turing, Kleene, Moore, ecc.).

V. Il circuito autoreferenziale gödeliano.

I teoremi di Gödel si basano sul fatto che l'aritmetica è un algoritmo universale, cioè un algoritmo capace di interpretare e tradurre nel suo proprio formalismo qualunque altro algoritmo, e che le dimostrazioni logiche possono essere interpretate e tradotte in calcoli aritmetici. L'idea nasce dal fatto che le elaborazioni logiche concrete, ad esempio le dimostrazioni dei teoremi di un sistema assiomatico, non sono altro, in ultima analisi, che produzioni di dati a partire da altri dati, esattamente come lo sono i calcoli aritmetici. In altri termini, le operazioni della logica booleana, di cui consiste una dimostrazione, non sono altro che un algoritmo applicato all'elaborazione di dati consistenti di proposizioni ben formate. Ora, l'aritmetica, giacché è un algoritmo universale, è capace di enumerare in modo ordinato tutti gli assiomi e le proposizioni ben formate di una teoria logica. Può anche rappresentare le deduzioni dei teoremi dagli assiomi di una qualsiasi teoria logica (ad esempio la geometria euclidea o la meccanica razionale), o da altri teoremi della stessa teoria, nella forma di operazioni aritmetiche condotte sui numeri che codificano gli assiomi o questi altri teoremi. La danza delle complessificazioni e delle semplificazioni che intervengono in una dimostrazione logica si tradurrà, in corrispondenza, in una danza di allungamenti e raccorciamenti di espressioni numeriche. Procedendo in questo modo l'intero sistema dei ragionamenti logici ammissibili dalla teoria così codificata sarà rappresentato da un sistema di operazioni aritmetiche.

Il passaggio chiave del procedimento gödeliano consiste in questo: *se si applica la codificazione aritmetica alla stessa teoria assiomatica dell'aritmetica si genera un circuito autoreferenziale dotato di straordinarie implicazioni e conseguenze*.

In questo modo, l'algoritmo aritmetico è messo nelle condizioni d'interpretare se stesso e di dimostrare, nella forma di particolari calcoli aritmetici, le sue stesse proprietà logicamente dimostrabili. Così ad esempio il teorema "esistono infiniti numeri primi", proposizione che, si badi bene, non è verificabile analizzando caso per caso quali numeri siano primi, sarà codificata da un numero particolare. La dimostrazione del teorema si presenterà allora nella forma di una successione di calcoli aritmetici che mettono in relazione tale numero con quelli che codificano gli assiomi. Insomma in questo modo l'aritmetica, intesa come teoria logica, vale a dire non come algoritmo ma come sistema assiomatico, riesce a codificare in termini di calcoli aritmetici *particolari* le proprietà *generali* del calcolo aritmetico. Questo può essere fatto per tutto ciò che è dimostrabile nell'ambito della teoria aritmetica. In modo più generale possiamo affermare che mediante la *gödelizzazione* il *linguaggio* aritmetico riesce a inglobare tutte le proposizioni *logicamente dimostrabili* del suo metalinguaggio.

VI. L'esplosione della complessità indescrivibile.

Il punto cruciale dell'indagine gödeliana si raggiunge quando ci si rende conto che un metalinguaggio contiene necessariamente più proposizioni del suo linguaggio oggetto, allo stesso modo come il numero dei sottoinsiemi di un insieme contiene più elementi dell'insieme stesso. Questo fatto può essere formulato con maggiore precisione affermando che la *cardinalità* delle proposizioni del metalinguaggio è superiore a quella del linguaggio. Una breve digressione potrà servire a chiarire le idee su questo aspetto della questione.

Due insiemi possiedono la stessa cardinalità se possono essere mappati uno sull'altro. Nel caso degli insiemi finiti la cardinalità equivale al numero di elementi. Ma la nozione di cardinalità si estende anche agli insiemi infiniti. Un insieme infinito può essere sempre mappato in una sua parte propria, tanto che un insieme può definirsi infinito se gode di questa proprietà. Ad esempio tutti i numeri interi possono essere mappati nei soli numeri pari mediante una semplice moltiplicazione per 2. Si può anche dimostrare, per esempio, che tutti i punti di un quadrato possono essere mappati su uno dei suoi lati e viceversa. Ma in molti altri casi queste mappe tra insiemi infiniti non possono farsi. Così, ad esempio, come ha insegnato Cantor, mentre i numeri interi possono essere mappati sui razionali e viceversa, non si può costruire una corrispondenza biunivoca tra i punti di un segmento e l'insieme dei numeri naturali giacché la cardinalità del continuo è superiore a quella del discreto.

Ora, tornando all'argomento iniziale, il linguaggio aritmetico ha la cardinalità del discreto, dato che la totalità dei calcoli aritmetici possibili può essere ordinata ed enumerata. Ma il suo metalinguaggio, avendo come universo del discorso le classi dei calcoli aritmetici possibili, ha la cardinalità del continuo. Stando così le cose, si comprende che devono esistere infinite proposizioni del metalinguaggio che non sono interpretabili dal linguaggio. Bisogna anzi aggiungere che si tratta di un'infinità infinitamente preponderante! Dato che la costruzione gödeliana mappa la parte logicamente dimostrabile del metalinguaggio aritmetico nel linguaggio aritmetico, si deduce che il metalinguaggio aritmetico possiede infinite *proposizioni indecidibili*, o teoremi indimostrabili, che sono indimostrabili perché altrimenti le dimostrazioni avrebbero una *complessità indescrivibile*. Il grande logico austriaco ha concretamente mostrato come si possa costruire una proposizione indecidibile.

VII. L'equivalenza dei punti di vista di Gödel e di Turing.

Originariamente l'analisi di Gödel riguardava il rapporto tra la struttura logica e la potenza algoritmica della matematica. Ciò rifletteva in modo naturale gli interessi maturati nel dibattito scientifico del primo '900. I risultati di Gödel furono reinterpretati da Turing come problema del rapporto tra la struttura e il comportamento di una macchina calcolatrice. La teoria delle macchine di Turing riduce all'osso la sostanza di questo problema.

La macchina di Turing universale, benché abbia una struttura, tutto sommato, abbastanza semplice, può fare tutto quello che può essere fatto dai calcolatori più complessi. Questa possibilità dipende in definitiva dal fatto che la struttura della macchina più complessa viene codificata come un particolare insieme di dati della macchina più semplice. Grazie a questa riduzione, l'intera teoria dei processi di calcolo si riconduce allo studio delle possibilità computazionali di questo congegno di complessità matematica descrivibile in poche pagine. In pratica, l'universalità della macchina di Turing si identifica con la sua capacità di eseguire qualunque calcolo aritmetico esattamente come l'universalità dell'algoritmo aritmetico si identifica nella possibilità computazionali della macchina di Turing universale. L'analogo del teorema di indecidibilità di Gödel si ha nel fatto che una macchina di Turing universale è sempre in grado di simulare un'altra, ma, in generale, non di prevedere se un processo di calcolo effettuato da quest'altra macchina si arresterà dopo un numero finito di passi (problema dell'*halting*).

VIII. Universalità costruttiva e automi capaci di autoriprodursi.

La teoria degli automi di von Neumann ricalca lo stesso schema concettuale. Per certi aspetti, essa rappresenta una naturale estensione del punto di vista di Turing. Un *automa costruttore* è una macchina capace di usare oggetti reperibili nell'ambiente per produrre altri oggetti da riporre nel medesimo ambiente. A tale scopo esso deve disporre di un elenco d'istruzioni e possedere un repertorio sufficientemente ricco di sensori e strumenti di lavoro per eseguire tutte le necessarie operazioni di cernita e assemblaggio sulla base di dette istruzioni; perciò, in generale, deve anche essere capace di elaborare informazione. Un automa di complessità strutturale insufficiente non riuscirà a produrre altro che oggetti di complessità inferiore alla sua. Ma se l'automa è dotato di organi capaci di svolgere un numero sufficiente di funzioni elementari, ed è inoltre sufficientemente bene organizzato, allora esso diviene *universale*. Vale a dire che, leggendo un opportuno programma d'istruzioni, esso diventa capace di produrre qualsiasi oggetto, anche automi di complessità uguale o superiore alla sua; dunque anche una copia di se stesso. A questo punto è chiaro che la cellula vivente ha tutti i requisiti di un automa universale!

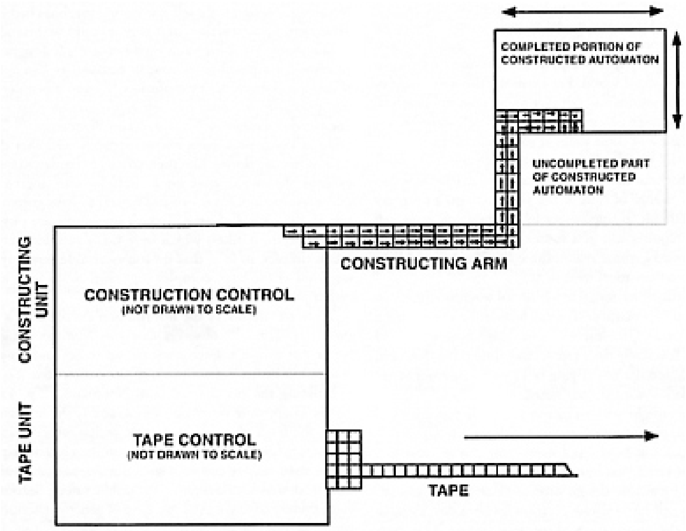

Fig.8. Schema semplificato del costruttore universale descritto da von Neumann in *The Theory of Self-Reproducing Automata*.

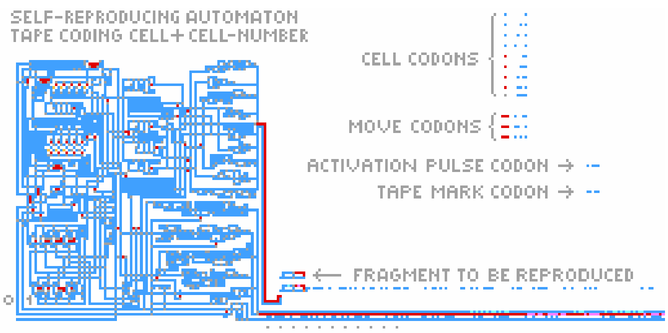

Fig.9. Costruttore universale planare a 32 stati capace di autoriprodursi indefinitamente, di tipo simile a quelli studiati da von Neumann . Il programma dell'automa, implementato dall'autore, è disponibile in rete (Nobili, 2004). La prima versione, implementata da Nobili e Pesavento nel 1996, era capace di riprodursi una volta sola.

Von Neumann - richiamandosi al teorema di indecidibilità di Gödel - afferma che l'automa non potrà "auto-copiarsi" esaminando direttamente la propria struttura interna mediante sensori o altri ipotetici apparati d'introspezione, poiché nessuna automisurazione o auto-osservazione può essere completa. Infatti, nessun apparato di misura può compiere una misura dentro se stesso. Affinché l'autoriproduzione abbia luogo è essenziale che l'automa costruttore disponga di una descrizione della sua struttura interna e di un apparato di decodificazione che gli permetta di dedurre le procedure costruttive o, equivalentemente, di un elenco completo delle istruzioni riguardanti le modalità di costruzione e assemblaggio delle sue parti. Tale è appunto il codice genetico degli organismi viventi.

Si ha in ciò l'analogo della codificazione di un sistema assiomatico nei dati di un algoritmo universale, senza la quale l'interpretazione autoreferenziale non potrebbe effettuarsi. È notevole il fatto che von Neumann giunse ad affermare tutto questo circa tre anni prima della scoperta del DNA (Watson e Crick, 1953), suggerendo, tra l'altro, sulla base di argomenti relativi a condizioni di funzionalità ottimale, che l'informazione genetica sia codificata in sequenze lineari di dati (come i simboli sui nastri delle macchine di Turing). Il sommo matematico ha fornito l'esempio di un semplice automa capace di autoriprodursi costituito di cellette quadrate uguali, ciascuna dotata di 29 stati interni, che possono saltare da uno stato all'altro secondo certe regole di transizione e in modi dipendenti dagli stati delle prime vicine.

Il punto cruciale di questa profonda visione, che può ritenersi l'istanza fondante della biologia teorica, è il ruolo svolto dalla distinzione tra la *complessità strutturale* e la *complessità comportamentale* di un sistema dotato di un numero finito di stati. Sotto un certo livello di complessità strutturale il sistema non riesce ad evolvere in modi molto complicati. In queste circostanze, le possibilità di calcolo di un algoritmo, i comportamenti di una macchina calcolatrice o di un automa costruttore appaiono prevedibili e descrivibili in termini finiti. Ciò comporta che la potenza interpretativa dell'algoritmo, quella di calcolo del calcolatore o la capacità produttiva dell'automa sono limitate. Pertanto l'universalità algoritmica, quella computazionale e quella costruttiva, quindi la gödelizzazione e l'autoriproduzione, sono impossibili.

Ma se il sistema possiede una struttura sufficientemente ordinata e articolata, e costruita in modo da permettere forme di ricorsività abbastanza complicate, si produce un fatto sconvolgente: la varietà dei comportamenti qualitativamente diversi esplode esponenzialmente verso il limite ideale del continuo, e accade che la maggioranza di questi comportamenti non siano prevedibili e descrivibili sulla base della conoscenza della struttura che li genera. L'unico modo di conoscerli è quello di osservarli nel loro effettivo svolgimento. In queste circostanze la potenza interpretativa dell'algoritmo, la capacità di simulazione del calcolatore o quella produttiva dell'automa diventano universali.

IX. Lingue, dialetti, universalità interpretativa e potenza creativa.

In analogia con le nozione finora introdotte, possiamo dire che un linguaggio è universale se riesce a tradurre bene i testi di qualunque altro linguaggio, nonostante le inevitabili differenze di struttura grammaticale, sintattica e di campo semantico. Le grandi lingue nazionali, come l'italiano, il francese, l'inglese, il tedesco ecc., sono linguaggi universali. I dialetti invece, a causa della loro povertà grammaticale e sintattica, non riescono a tradurre bene i significati dei testi più sofisticati di altre lingue (ad esempio, nonostante la mia confidenza col dialetto veneto, non riesco assolutamente ad immaginare come la *Fenomenologia dello Spirito* di Hegel potrebbe essere tradotta in questo dialetto conservando i significati originali e senza produrre effetti di comicità irresistibile).

Ciò che marca la differenza tra una lingua e un dialetto è che la prima, proprio perché è un linguaggio universale, riesce a descrivere sé stessa, ad essere, cioè, autoreferenziale, in altri termini di funzionare in modo coerente sia come linguaggio sia come metalinguaggio. Le lingue universali non sono nate spontaneamente dalla pratica della comunicazione, ma si sono evolute e raffinate grazie al lavoro paziente e accurato di pensatori, scrittori, filosofi e scienziati, che si sono trovati nelle condizioni di dover escogitare i mezzi espressivi più adeguati per esprimere pensieri molto complessi e culturalmente evoluti. Ora, come un algoritmo universale è in grado non solo di tradurre i testi di altre lingue ma anche di produrre strutture di dati di complessità illimitata, allo stesso modo un linguaggio universale riesce ad organizzare qualsiasi forma di comunicazione possibile e immaginabile; di esprimere concetti letterari, filosofici e scientifici di complessità arbitrariamente grande, favorendo in tal modo le capacità creative del pensiero. Si spiega così l'illimitata potenza espressiva che le lingue nazionali hanno rispetto ai gerghi e ai dialetti.

Un discorso simile può farsi anche a proposito della creatività artistica. La differenza sostanziale tra un artista dilettante e uno veramente creativo è che il primo si limita a copiare o a imitare, mentre il secondo è dotato di capacità espressive e interpretative universali che gli permettono di esprimere liberamente tutte le sensazioni, le emozioni e le idee che gli passano per la mente. Sia nelle arti figurative che nella musica, i mezzi espressivi di un artista e la sua tecnica costituiscono un linguaggio che, nel caso di arte autentica, è praticamente intraducibile sia in lingua scritta sia in qualunque altra forma espressiva. La grammatica e la sintassi di questo linguaggio sono creazioni personali dell'artista, e la loro struttura profonda traspare solo dalla coerenza stilistica delle sue opere. Anche in questo caso la ricorsività del processo di elaborazione è essenziale per la creatività. L'artista prova e riprova, un pò a caso e un pò seguendo alcune idee, selezionando i risultati secondo il suo gusto estetico, fino ad ottenere qualcosa che gli appare autoconsistente e dotato di capacità di comunicazione universale.

1.3. Il problema dell'autocoscienza

Un'importante estensione della teoria delle macchine ricorsive investe direttamente il problema dell'autocoscienza. Molti autori hanno osservato che l'autocoscienza non può nascere da un semplice processo d'*introspezione*. Ciò è perfettamente comprensibile poiché la tesi dell'introspezione darebbe luogo a regressioni all'infinito del genere: *io* sono consapevole di guardare il tavolo; → *io* sono consapevole di essere consapevole di guardare il tavolo; → *io* sono consapevole di essere consapevole di essere consapevole di guardare un tavolo; ecc. Questa situazione è ben rappresentata dall'immagine, riportata in *Anelli nell'Io* di Hofstadter, che rappresenta il personaggio di un fumetto di Bushmiller, di nome Sluggo, che riflette su sé stesso (Fig.10).

Fig.10. Sluggo.

La questione è per alcuni aspetti analoga a quella della macchina costruttrice universale, che non può copiare sé stessa mediante un semplice procedimento di auto-ispezione ma lo può fare con l'ausilio di un programma che ne codifica la struttura. Questo suggerisce che anche la "macchina della mente" sia dotata di un programma che ne codifica la struttura. In questo modo, quella entità che nell'esempio su riportato si presenta come "*Io che guardo il tavolo*" sarebbe in realtà l'immagine costruita da un programma capace di generare l'immagine di un personaggio chiamato *Io* che guarda l'immagine di un tavolo. Incidentalmente, questo si accorda abbastanza bene con ciò che è testimoniato da persone soggette a stati di dissociazione mentale, le quali affermano di avere talvolta la sensazione di osservare sé stesse dall'esterno.

Con ciò si vuole asserire che ogni volta che noi diciamo *io*, in realtà non ci riferiamo, come siamo portati a credere, a noi stessi in carne ossa e nervi, né all'attività pensante in atto del nostro cervello, ma ad *una rappresentazione di noi stessi posti in relazione con la rappresentazione di altre cose*. Rappresentazione che, in omaggio ad Hofstadter, possiamo chiamare il *simbolo del sé*.

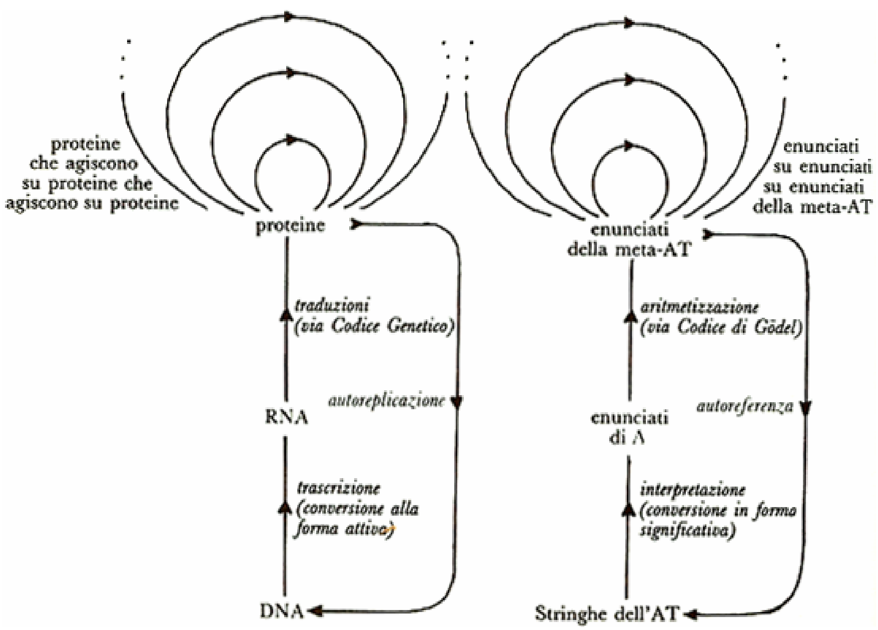

Fig.11. Come Douglas Hofstadter vede la relazione tra l'autoriproduzione biologica l'autoreferenza logico-aritmetica (Da Hofstadter 1985).

Il merito di Hofstadter va riconosciuto perché questo brillante autore introdusse per primo, nel suo libro *Gödel, Escher e Bach*, l'idea che il processo mentale che genera l'autocoscienza assomigli al procedimento usato da Gödel nelle sue ricerche sulle possibilità di calcolo dell'aritmetica.

Essa ricalca in buona sostanza lo schema della corrispondenza, indicata dallo stesso Hofstadter (Fig.11), tra i processi che stanno alla base della capacità di autoriprodursi degli esseri viventi e quelli che stanno alla base delle procedure attraverso con le quali Gödel mise in evidenza la capacità autoreferenziale dell'aritmetica.

Nella visione di Hofstadter, il simbolo del sé svolge il ruolo della teoria assiomatica dell'aritmetica, che, in un certo senso, codifica la struttura dell'algoritmo aritmetico, o quello del programma genetico dei viventi che codifica la struttura dell'organismo.

In *Anelli nell'Io* (2004) Hofstadter precisa che il simbolo del sé é generato da una processo ricorsivo dotato di programma residente nel cervello. In questo stesso libro, Hofstadter arriva anche a dichiarare che i processi generatori del simbolo del sé possono persino generare, sia pure in forma larvale, i simboli del sé di altri esseri umani coi quali il soggetto pensante principale abbia sufficientemente interagito e comunicato nel corso della sua vita.

Dipanando l'argomento da metafore, doppi sensi, giochi di parole e dallo stile fantasioso da *Alice nel Paese delle Meraviglie* che avviluppano lo spirito del libro, essa può essere brevemente descritta in questo modo: un cervello pensante è la sede di processi di produzione di certi stati di eccitazione della rete nervosa, che l'autore definisce *simboli attivi*. Questi sono così definiti perché, a differenza dei simboli passivi che intervengono nei formalismi matematici, non ricevono i loro significati da relazioni di corrispondenza poste da un soggetto esterno, ma sono essi stessi capaci di "attivare" e reclutare relazioni con oggetti esterni. Per mantenere l'analogia coi casi precedenti bisogna assumere che la "macchina mentale" sia capace di trasformare i simboli attivi in altri simboli attivi, secondo procedure che vengono istruite in qualche modo non conosciuto. Naturalmente, non deve esistere alcun apparato supervisore, nessun *homunculus* cosciente nascosto che legga e interpreti un elenco di istruzioni al fine di far eseguire al cervello tali operazioni (Fig.12) altrimenti si produrrebbe una regressione all'infinito che non spiega nulla. L'attività di produzione e riproduzione simbolica di cui consiste ciò che si usa chiamare *mente* deve generarsi e sostenersi autonomamente per interazione causale dei simboli attivi; simboli che si attivano e si disattivano nella struttura cerebrale secondo una specifica dinamica. In analogia coi i concetti illustrati nei paragrafi precedenti, aggiungiamo la seguente definizione: la mente A è capace di *interpretare* la mente B se A è capace di stabilire una corrispondenza tra ogni produzione simbolica di B e una delle sue.

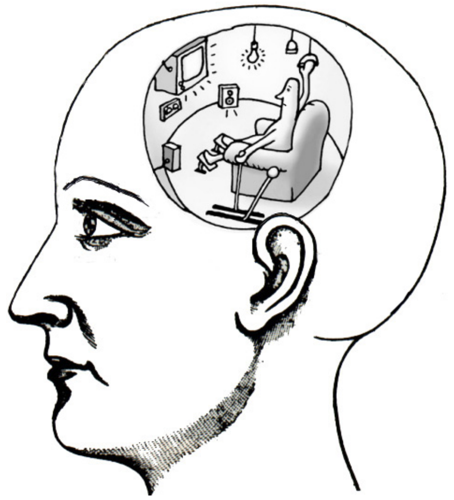

Fig. 12. Homunculus annidato nel cervello.

Affinché ciò possa avere luogo, A avrà bisogno di mezzi atti ad acquisire e decodificare i simboli e le produzioni simboliche di B; dovrà pertanto possedere una capacità di rappresentazione interna coordinata ad apparati d'interazione col mondo esterno. Una mente che abbia una complessità strutturale insufficiente, ad esempio quella di una scimmia, non può interpretare altro che produzioni simboliche di complessità limitata. Ma se la sua struttura è abbastanza complessa ed articolata, e la sua memoria è abbastanza capace, essa diventa una mente *universale*. Purché sia opportunamente istruita, essa potrà interpretare ogni altro genere di produzione simbolica: quella di una macchina calcolatrice, di un animale, di altre menti dotate di qualsiasi grado di complessità. E' evidente che, in questo senso, *la mente umana è universale*!

Hofstadter non spiega come in un cervello animale possa formarsi una mente universale, giacché nel libro citato, o in un altri del medesimo autore, la questione dell'universalità risulta implicita solamente nell'idea della corrispondenza coi teoremi di Gödel. Tuttavia, a parte considerazioni di carattere filogenetico, possiamo ragionevolmente ipotizzare che tale capacità si organizzi fin dalle primissime fasi dell'età evolutiva attraverso la comunicazione espressiva degli stati emotivi, l'esercizio della loro simulazione, l'apprendimento emulativo e l'assimilazione dei codici di comportamento. Tutto questo corrisponde perfettamente a quanto Antonio Damasio (1994, 2000, 2004) descrive nei suoi stupendi libri a proposito della base neurologica delle emozioni. Ma, naturalmente, lo sviluppo di queste facoltà non basta a produrre una capacità d'interpretazione universale.

Nemmeno riguardo ai meccanismi dell'autocoscienza lo studioso americano fornisce spiegazioni adeguate all'idea che avanza. Possiamo tentare di completare in questa sede la sua visione spingendo fino in fondo l'analogia con la costruzione gödeliana e incorporando in questa i concetti chiave della teoria degli automi di von Neumann. La questione sembra riguardare il rapporto tra due fondamentali facoltà della mente umana: l'*immaginazione*, identificabile con l'attività di produzione simbolica descritta sopra, e il *linguaggio*, inteso in senso generalizzato, cioè come la totalità dei mezzi di espressione degli stati mentali e di comunicazione interpersonale. Sembra ragionevole, infatti, che si possa stabilire una corrispondenza tra queste due facoltà e i modi di porsi della matematica discussi nel punto IV del paragrafo 1.2: quello algoritmico e quello logico. La gödelizzazione del rapporto tra l'immaginazione e il linguaggio potrebbe compiersi nel seguente modo.

L'attività pensante, che dal punto di vista neurologico-strutturale sembra consistere di un sistema di aree e centri cerebrali interagenti, capaci di reclutare ed evocare associativamente i contenuti della memoria sotto forma di stati di eccitazione collettiva dei neuroni, dal punto di vista neurologico-comportamentale si presenta come una produzione indescrivibilmente complessa di stati di eccitazione neuronale, che si presentano soggettivamente come elaborazioni di fantasie specifiche particolari. Indipendentemente dalle relazioni che i dati di memoria possano avere avuto in origine con la percezione, l'elaborazione di una fantasia ha proprietà formali simili a quelle di una produzione algoritmica. Essendo una produzione ricorsiva di stati mentali di complessità indescrivibile, essa non può essere riassunta da uno stato mentale particolare, ma può essere interpretata e compresa da un'altra mente, possiamo anche dire simulata, solo tramite la corrispondenza con una produzione analoga di quest'altra mente. Naturalmente questo può avvenire solo se ciascuna delle due menti possiede un proprio apparato per la manifestazione simbolica delle fantasie e un sistema di codificazione simile a quello dell'altra, dunque un linguaggio comune. Nessuna mente è in grado di percepire direttamente le fantasie di un'altra mente, ma può ricostruirle nella forma di fantasie proprie se i messaggi ricevuti sono sufficientemente coerenti e dettagliati. Una mente che sia in grado di immaginare le fantasie di qualsiasi altra mente e che riesca a comunicarle in modo adeguato potrà definirsi universale.

Sulla base di queste considerazioni possiamo identificare il significato di un messaggio, non tanto con un insieme di dati o stati di eccitazione neuronale che lo rappresentano nel cervello, quanto piuttosto come il processo ricorsivo indescrivibilmente complesso che esso attiva nella macchina cerebrale.

Ora, come un algoritmo universale è dotato di capacità autoreferenziale, così una mente universale può interpretare e comprendere anche sé stessa, ossia esemplificare nella propria immaginazione mediante i suoi stessi mezzi di comunicazione, anche le fantasie da essa stessa liberamente o casualmente generate, indipendentemente dal fatto che queste fantasie siano derivate dai dati della memoria. Affinché ciò possa funzionare, l'immaginazione e la comunicazione devono consistere di due attività indipendenti che entrano in relazione tra esse nei processi interpretativi del pensiero. Nel loro insieme, esse impartiscono all'attività mentale un carattere dualistico simile a quello che si ritrova nella distinzione tra l'algoritmo aritmetico e la teoria aritmetica. L'autocoscienza, nella sua essenza, dovrebbe consistere proprio nella gödelizzazione dell'attività pensante, precisamente come un procedimento autoriflessivo capace di produrre le interpretazioni immaginative delle produzioni linguistiche, a loro volta indotte dall'attività immaginativa stessa.

Perseverando nell'analogia, e rifacendosi alle tesi di von Neumann, dobbiamo ipotizzare che una mente universale non possa giungere all'autocoscienza semplicemente "percependo" o "guardando dentro se stessa", poiché l'autopercezione diretta deve ritenersi tanto impossibile quanto l'autointerpretazione diretta della teoria aritmetica senza il passaggio attraverso la codificazione algoritmica delle sue proposizioni, o l'autoriproduzione degli organismi viventi senza la codificazione dell'informazione genetica in un DNA.

Ma dobbiamo assumere che, come nel caso delle macchine di Turing e di von Neumann l'universalità interpretativa dipende dalla possibilità di riprodurre e simulare al proprio interno i processi di altre macchine, allo stesso modo l'universalità di una mente dipenda dalla possibilità di rappresentare gli stereotipi comportamentali di altre menti, come in una sorta di teatro delle marionette. Come una macchina di von Neumann dotata di universalità costruttiva è per questa stessa ragione capace di autoriprodursi, così una mente dotata di universalità interpretativa è capace di "autointerpretarsi". Per fare ciò essa deve utilizzare proprio quelle risorse che le permettono di interpretare le altre menti.

Dovremmo perciò ipotizzare che in una mente universale esista uno speciale sistema simbolico capace di rappresentare ed esemplificare immaginativamente anche la struttura formale della propria attività pensante, le relazioni tra le parti di questa struttura e le funzioni dei suoi meccanismi interni. Generalizzando la nozione di "simbolo del sé" introdotta da Hofstadter, possiamo chiamarlo *sistema simbolico del sé*.

In un certo senso, quest'ipotesi non fa altro che ripresentare in altra forma la tesi dell'homunculus annidato nel cervello. Solo che ora questo strano personaggio non deve essere immaginato come una struttura neuronale collocata in una parte speciale del nostro cervello e nemmeno essere identificato con gli *homunculi* somato-sensoriale e motorio, che si usano dipingere sulle mappe della corteccia cerebrale per indicare dove risiedono i neuroni che corrispondono alle diverse parti del corpo (Fig.13). Le strutture somato-sensoriali e motorie sono un aggregato di aree specializzate costantemente impegnate a codificare e veicolare i flussi d'informazione reciprocamente indipendenti tra cervello e soma. L'homunculus di cui si ha bisogno deve essere inteso come un'entità simbolica complessa che rappresenta in modo integrato lo schema di ciò che accade nel cervello, non i processi mentali effettivi nella loro complessità. Allo stesso modo, gli assiomi dell'aritmetica teorica non rappresentano tutto ciò che l'algoritmo aritmetico può generare, e la doppia elica del DNA non rappresenta ciò che accade nell'organismo vivente.

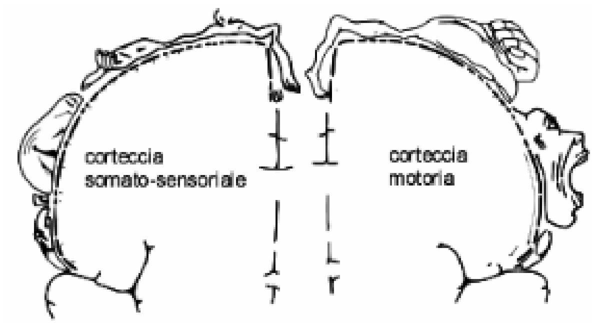

Fig. 13. Homunculi somato-sensoriale e motorio della corteccia cerebrale.

Il grande sogno dell'IA, la creazione dell'autocoscienza artificiale, sembra dunque condizionato dalla possibilità di costruire macchine elaboratrici di informazione capaci di comunicare linguisticamente con gli esseri umani, interpretarne il pensiero, introiettarne lo schema strutturale, fino ad elaborare un proprio sistema simbolico del sé e a diventare in questo modo anche capaci d'autointerpretazione. Può la moderna tecnologia dei calcolatori arrivare a tanto?

1.4. La macchina introspettiva di Konolige

La capacità autoreferenziale dell'aritmetica e quella autoriproduttiva del costruttore universale suggeriscono che una macchina di Turing universale possa funzionare in modo autoreferenziale, vale a dire, essere messa in condizioni di analizzare, descrivere e comunicare i calcoli che essa stessa ha effettuato o sta effettuando e le procedure che ha usato o sta usando. Una macchina calcolatrice dotata di questa capacità potrebbe essere definita *introspettiva*, sebbene questa denominazione sia considerata impropria da taluni autori. Essa fornirebbe un modello della capacità autoriflessiva del pensiero umano e una sua eventuale implementazione equivarrebbe alla creazione del pensiero artificiale. Negli anni '80 vari ricercatori hanno avviato alcune ricerche teoriche volte allo sviluppo di una teoria delle macchine introspettive (Bartlett, 1992).

In un articolo del 1985 Kurt Konolige definisce la *macchina introspettiva* come un sistema dotato di un *sottosistema di credenze*. Secondo la sua definizione, un sottosistema di credenze è una struttura computazionale all'interno di un agente artificiale capace di rappresentare un numero finito di credenze riguardanti i fatti del mondo. Questo sottosistema è in grado di accettare un quesito posto da un agente esterno, e tentare di verificare se la risposta al quesito può essere derivata dal suo sistema di credenze. Il funzionamento della macchina si basa sulla scomposizione di un quesito complesso in quesiti più semplici e sul successivo tentativo di rispondere ai sottoquesiti.

Konolige assume che la macchina introspettiva possieda per ogni sua credenza *C*, che egli definisce *non doxastica*, anche la credenza "*credo che C sia una mia credenza*", che egli definisce *doxastica*. Questo semplicemente significa che il linguaggio doxastico è il metalinguaggio di quello non doxastico. L'analogia coi casi precedentemente considerati è palese. Chiaramente, una condizione necessaria affinché la macchina sia effettivamente capace d'introspezione è che l'algoritmo che elabora i quesiti non doxastici sia universale. Il teorema d'indecidibilità di Gödel assicura allora che esistono infiniti quesiti ai quali la macchina non sa rispondere.

Konolige non lo dice, ma un'altra condizione perché la macchina possieda effettivamente la capacità introspettiva è che essa sia capace di porre autonomamente a se stessa tutti i quesiti che possono essere posti da un agente esterno. Può una macchina di Turing essere programmata per fare tutto questo? La risposta è negativa. La capacità di generare autonomamente i quesiti che possono essere imprevedibilmente posti da agenti esterni contrasta col determinismo della macchina di Turing. Questa capacità potrebbe essere supplita da un dispositivo capace di porre quesiti casuali "ben formati" sulla base di un'opportuna grammatica generativa. La difficoltà sta nel fatto che nessuna macchina deterministica è in grado di produrre una sequenza casuale di simboli.

La domanda deve dunque essere riformulata nel seguente modo: può una macchina di Turing universale, magari equipaggiata con una sorgente di quesiti casuali, essere programmata per funzionare come una macchina introspettiva? *In teoria sì*, perché la macchina di Turing universale possiede tutte le condizioni per la sua gödelizzazione. *In pratica no*, perché i tempi di calcolo dei procedimenti ricorsivi attraverso i quali essa può tentare di rispondere ai quesiti, costruendo e analizzando classi di credenze, crescono in generale esponenzialmente col numero delle credenze. Questo dipende dal fatto che la cardinalità delle classi di credenze è superiore a quella delle credenze. Il collo di bottiglia che impedisce nella pratica il processo introspettivo sta nel fatto che le macchine di Turing, e con esse ogni calcolatore di tipo ordinario, sono *sequenziali*, nel senso che esse sono costruite in modo da eseguire un numero relativamente piccolo di operazioni a ogni ciclo regolato dal loro orologio interno. La coscienza di un evento mentale ci servirebbe assai poco, e forse sarebbe un intralcio, se invece di giungere con un ritardo di circa mezzo secondo dopo che un'intenzione è stata elaborata inconsapevolmente dal cervello, come gli esperimenti di Libet e collaboratori (1979) hanno dimostrato, impiegasse un tempo maggiore per produrre i suoi effetti sul comportamento del soggetto pensante e sulla comunicazione con altri soggetti pensanti. La sua formazione richiede che la macchina cerebrale sia molto potente, oltre che ben organizzata, per generare in tempi utili i processi introspettivi. Forse per questa ragione essa è comparsa tardi nell'evoluzione delle specie.

1.5. La potenza di calcolo del cervello

Le proprietà caratteristiche dei calcolatori ordinari e della macchina di Turing, che li rappresenta tutti in modo paradigmatico, sono la limitatezza numerica delle operazioni eseguibili nell'unità di tempo e serialità dei processi di calcolo. Ciò significa: *a*) che il tempo impiegato da queste macchine per effettuare un insieme di calcoli indipendenti non può essere minore della somma dei tempi necessari per effettuarli singolarmente; *b*) che in pratica sono possibili solo elaborazioni che richiedono tempi di calcolo che crescono come una potenza piccola del numero di dati iniziali.

Un moderno calcolatore da tavolo può eseguire alcuni miliardi di operazioni per secondo in corrispondenza dei cambiamenti di stato del suo microprocessore. La velocità dei calcolatori dipende dal fatto che i segnali elettrici si propagano a velocità prossime a quelle della luce. La velocità dei segnali nervosi è dello stesso ordine di grandezza di quella del suono, essendo limitata dalla natura elettrochimica dei processi da cui dipende. Questa lentezza, a fronte del vantaggio evolutivo degli organismi più veloci, ha imposto ai sistemi nervosi degli animali di adeguarsi ad un principio di *massima parallelizzazione possibile a parità di funzioni*. Nel seguire questa via le specie animali sembrano non aver perso nulla delle possibilità offerte a priori dalla natura: al fine della lotta per l'esistenza, i processi di calcolo parallelo effettuati dai sistemi nervosi si rivelano assai più efficienti di quelli che potrebbero essere eseguiti, in modo più preciso ma più lento, da macchine calcolatrici seriali.

Fino a un decennio fa molti neurofisiologi sembravano concordi nell'affermare che le effettive unità di processamento dell'informazione nervosa nella corteccia cerebrale non sono i singoli neuroni, ma piccole popolazioni di neuroni eccitatori e inibitori comprendenti da alcune migliaia di neuroni. L'attività elettroencefalografica (EEG), che si registra principalmente nell'intervallo di frequenza 6-100 hertz, era interpretata come uno stato oscillatorio delle attività di sparo di questi neuroni dovuta alla retroazione ritardata dei neuroni inibitori su quelli eccitatori.

Sapere cosa sia esattamente l'attività EEG è rilevante perchè ciò permetterebbe di capire quali sono i meccanismi della dinamica neuronale e la loro funzione nei processi cognitivi. Una delle teorie più diffuse è quella del *caos dinamico* sostenuta da Walter Freeman (1999), che interpreta l'attività oscillatoria come la manifestazione di una dinamica in cui si alternano fasi di attività disordinata e processi di organizzazione spontanea tendenti verso precise configurazioni spaziotemporali (attrattori) ricche d'informazione.

Purtroppo questa teoria ha molti difetti. Innanzi tutto è noto da simulazioni computazionali che sistemi di oscillatori debolmente interagenti possono raggiungere uno stato di sincronizzazione in un tempo non inferiore a quello di alcuni cicli oscillatori. Così, ad esempio, un sistema di oscillatori neuronali operante alla frequenza di 80 hertz (attività gamma) impiegherebbe non meno di mezzo secondo per produrre una configurazione di attività neurologicamente significativa. Mentre si sa che i neuroni delle aree visive si sincronizzano in circa un millisecondo (Singer, 1999).

Inoltre, l'identificazione delle unità funzionali del cervello con piccole popolazioni di neuroni abbasserebbe di alcune migliaia di volte la complessità strutturale effettiva del cervello. Dato che il cervello umano contiene circa dieci miliardi di neuroni (10^{10}), si deduce che esso dovrebbe ospitare circa dieci milioni di tali piccole unità. Anche lavorando alla frequenza media di oscillazione di 40 hertz, il cervello potrebbe effettuare al più 400 milioni di operazioni al secondo. Si dedurrebbe che la potenza di calcolo del cervello umano è un centinaio di volte maggiore di quella di un moderno calcolatore da tavolo e circa uguale a quello di grossa centrale di calcolo.

Ma nel confronto tra le prestazioni del cervello e quelle del calcolatore, le differenze più serie sono quelle che riguardano la capacità di memoria e il rapporto tra il tempo di accesso alla memoria e i tempi dei cicli di calcolo. Queste differenze sono enormemente più sfavorevoli per il calcolatore. La capacità di memoria di una sinapsi neuronale si può stimare di circa 1 byte (8 bit). Dato che ogni neurone possiede mediamente diecimila sinapsi, possiamo stimare che la capacità di memoria di un neurone sia di circa 10 kilobyte e quella del circuito talamo-corticale sia di alcune decine di terabyte (migliaia di miliardi di byte). Probabilmente molto di più se anche altre cellule del cervello (in particolare gli astrociti) possiedono meccanismi di memoria. Circa la velocità di accesso alla memoria, quella di una sinapsi si può stimare 1-2 millisecondi (pari al tempo di attivazione di una sinapsi), che è una frazione trascurabile del tempo medio di un ciclo di un circuito riverberante neuronale.

Quanto alla memoria esterna dei calcolatori, bisogna considerare che oggi sono in commercio dischi fissi che hanno una capacità di memoria di alcuni terabyte (mille miliardi di byte), una quantità confrontabile con quella del cervello. Purtroppo il tempo di accesso ai dati memorizzati in un disco fisso è di alcuni millisecondi. Ciò ritarda in modo enorme la velocità di calcolo effettiva dei microprocessori. Infatti, per eseguire un'operazione di calcolo utilizzando dati memorizzati nel disco fisso bisogna prima leggere i dati dal disco e poi scrivere i risultati sul disco (*swapping*). Credo che tutti coloro che hanno un calcolatore con poca memoria RAM (*random access memory*) sappiano quanto sono noiose queste operazioni di *swapping*.

Per contrasto, la memoria RAM, che funziona in modo totalmente elettronico, ha un tempo di accesso di circa 20 nanosecondi (circa due cicli macchina), che dunque non è trascurabile rispetto al ciclo di calcolo del microprocessore. Questo in pratica diminuisce il tempo effettivo del ciclo di calcolo di cinque volte (2+2+1). In vista dei futuri progressi della tecnologia, possiamo dire che se la memoria RAM fosse tanta quanta quella di un disco fisso, la potenza di calcolo effettiva di un calcolatore si avvicinerebbe molto a quella di un cervello umano. Purtroppo però le memorie RAM ora disponibili sono al massimo dell'ordine di grandezza di alcuni gigabytes (miliardi di byte).

Queste cifre ci danno la misura di quanto sia distante la possibilità di simulare con un calcolatore le funzioni del cervello umano. Esse spiegano inoltre perché il sistema visivo è capace di riconoscere un oggetto posto in un ambiente complesso in un decimo di secondo, mentre un calcolatore, per svolgere questa funzione con un grado di efficienza accettabile, impiegherebbe un tempo migliaia di volte maggiore. Bisogna tuttavia considerare che il progresso tecnologico sembra non conoscere limiti, cosicché non è azzardato pensare che in un futuro non tanto lontano sarà possibile disporre di sistemi di processamento d'informazione capaci di simulare in tempo reale i processi paralleli del cervello umano. E' pertanto legittimo chiedersi se si riuscirà a produrre un automa "autocosciente".

La differenza sostanziale tra il cervello e il computer sta dunque nell'enorme differenza di rapporto tra la quantità di dati che possono essere letti e riscritti in un ciclo di attività EEG del cervello e quella di un microprocessore: alcuni miliardi di byte nel primo caso, poche decine di byte nel secondo caso. Questa enorme differenza è dovuta al fatto che il calcolatore processa l'informazione in modo seriale, mentre il cervello la processa in modo parallelo.

1.6. Processi seriali e processi paralleli

C'è davvero una grande differenza di efficienza tra le macchina parallele e quelle seriali? E se c'è, come si spiega? Negli anni recenti l'elettronica e l'informatica hanno cercato di affrontare in modo sistematico, tanto nella teoria quanto nella pratica, il problema della costruzione di calcolatori paralleli. Le difficoltà che si sono incontrate testimoniano quanto l'informatica sia ancora lontana dalla possibilità di simulare la macchina della mente.

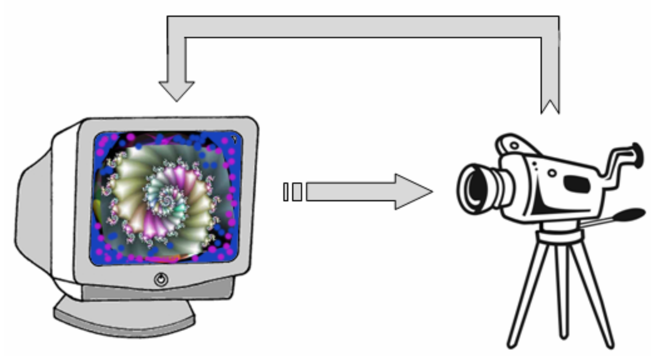

Fig.14. Processo ricorsivo video-televisivo. Il circuito genera immagini astratte complesse che evolvono in modo imprevedibile.

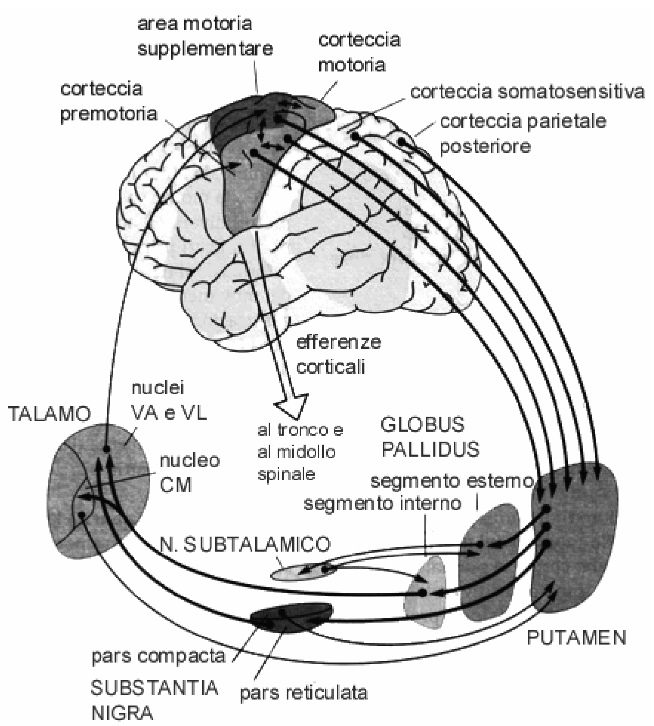

Fig.15. Il circuito talamo-corticale.

Le macchine seriali trasmettono e ricevono sequenze di dati attraverso poche linee di comunicazione. Tuttavia, quando operano in modo ricorsivo, esse sono capaci di produrre risultati di complessità indescrivibile. Il cervello lavora invece con flussi di dati molto intensi che sono trasmessi e ricevuti simultaneamente da apparati spazialmente estesi: gli apparati sensoriali visivo, tattile, muscolare, le aree della corteccia, i gangli basali nuclei encefalici ecc. e tutti gli organi che costituiscono il soma e sembra mancare di precisione sufficiente per poter funzionare in modo ricorsivo. Può una macchina seriale effettuare un processo di elaborazione di informazione equivalente a quello di una macchina parallela contando sulla sua maggiore velocità? Viceversa, può una macchina parallela effettuare processi ricorsivi di complessità equivalente a quella di una macchina seriale contando sulla sua maggiore complessità? Quale dei due tipi di macchine è in grado di fornire le prestazioni migliori per l'implementazione della capacità autoreferenziale del pensiero?

Con procedure di scansione veloce, un

insieme di dati presentati in modo sincronico può essere facilmente convertito in una sequenza di bit. Reciprocamente, una sequenza di bit può essere rapidamente trasformata in un insieme sincronico di dati. Le telecamere e gli schermi televisivi fanno proprio questo. Data la velocità dei processi di scansione elettronica, i calcolatori sono in grado di ricevere in ingresso e produrre in uscita flussi di dati paralleli simili a quelli che entrano ed escono nel cervello umano. In *Anelli nell'Io*, Hofstadter ha portato come esempio di processo ricorsivo parallelo, potenzialmente capace di generare immagini di grande complessità, il circuito illustrato in Fig.14. Si tratta di una telecamera disposta coassialmente di fronte ad un televisore che riproduce le immagini fornite dalla stessa telecamera in modo da conservarne le dimensioni. Nonostante il processo sia governato da dispositivi completamente seriali, si riesce a fornire un esempio di come potrebbero funzionare i processi ricorsivi del cervello. Ad esempio, si può ipotizzare che il circuito talamo-corticale illustrato in Fig.15 sia paragonabile al circuito televisivo su descritto, col talamo nel ruolo di telecamera e la corteccia in quello di televisore. Affinché una macchina seriale possa competere con una parallela, bisogna che riesca ad eseguire in sequenza un grandissimo numero di operazioni logico-aritmetiche su insiemi di dati man mano che questi arrivano in ingresso. Questo problema si presenta, ad esempio, nelle procedure di codificazione e decodificazione dei segnali video, che sono correntemente applicati a monte e a valle dei sistemi di trasmissione dei dati per aumentare la risoluzione delle immagini riprodotte dagli schermi. L'informazione analogica prodotta dalla telecamera è prima convertita in informazione digitale, sottoposta in questa forma a procedimenti di compressione (zippaggio), quindi trasmessa in formato compresso ai videoricevitori e infine decompressa e visualizzata dai display.

Tanto per dare un'idea del problema che si cerca di risolvere, supponiamo di utilizzare un calcolatore da tavolo di buona qualità per trasmettere con un piccolo ritardo di tempo un film che in formato non compresso occupa circa un megabyte (8 milioni di bit) di memoria per fotogramma alla velocità di 24 fotogrammi al secondo. Per comprimere nel modo più efficiente questo imponente flusso d'informazione bisogna analizzare le correlazioni che esistono tra tutte le parti di ogni fotogramma e tra alcuni fotogrammi successivi e ricodificare l'informazione in modo da ridurre drasticamente gli elementi d'informazione ridondanti. Questo processo richiede che grandi quantità di dati siano letti e scritti molto velocemente in una memoria temporanea di lavoro. Dato che il minimo tempo di accesso ai dati di un disco fisso è di alcuni millisecondi, non si può pensare di utilizzare dei dischi fissi per eseguire queste operazioni di lettura e scrittura. Per fortuna la memoria elettronica, o RAM (*random access memory*), è centomila volte più veloce, sebbene la sua capacità, fino ad alcuni miliardi di byte, migliaia di volte inferiore a quella di un disco fisso, essa permette di eseguire il lavoro di compressione e decompressione dell'informazione contenuta nel film all'incirca in tempo reale. Con questo tipo di memoria di lavoro il calcolatore seriale è in grado di simulare processi ricorsivi paralleli con flussi d'informazione comparabili a quelli dei cervelli di piccoli animali (ma migliaia di volte più piccoli di quelli del cervello umano). Si può pensare che tra alcuni anni i grandi calcolatori possano raggiungere la capacità e la velocità di elaborazione del cervello umano.

Insomma la differenza sostanziale tra una macchina seriale e una parallela come il cervello, se questo fosse capace di operare in modo ricorsivo, è che la prima deve eseguire un numero enorme di cicli di calcolo per ottenere quello che la seconda potrebbe fare in pochi cicli.

Una caratteristica distintiva dei calcolatori seriali è la netta separazione dell'hardware dal software, della macchina che esegue le operazioni dai dati e dai programmi contenuti nei sistemi di memoria. Questa separazione risponde al requisito di ottimizzazione dei processi seriali. Il modo di esecuzione di un processo seriale dipende sia dall'architettura del microprocessore, che è fissa, sia dai programmi di calcolo, che fanno parte dei dati immagazzinati nella memoria e variano da caso a caso. La differenza tra i programmi e i dati da processare sta solo nei modi con cui questi diversi tipi di dati sono utilizzati dal microprocessore.

L'informatica dei processi paralleli identifica, in pratica, il calcolatore parallelo con un certo numero di calcolatori seriali che si scambiano messaggi (Codenotti e Leoncini, 1990). Il nostro interesse è rivolto invece ai sistemi di elaborazione di dati in cui la parallelizzazione è massima compatibilmente con la condizione della massima efficienza. E' plausibile che i principi teorici che devono esser posti alla base del calcolo parallelo siano diversi da quelli su cui si è basata la teoria del calcolo seriale. Cercheremo di impostare il problema senza pretendere di arrivare ad una soluzione generale. Nel seguito intenderemo per calcolatore parallelo un calcolatore i cui processi di calcolo sono massicciamente parallelizzati. Dato che per ora calcolatori di questo tipo non esistono, il meglio che si può fare e di cercare di capire quali caratteristiche dovrebbero avere.

Affinché una macchina parallela sia capace di competere con una seriale nello svolgimento di processi ricorsivi bisogna che riesca a mantenere con grande precisione lo stato di sincronizzazione dei flussi d'informazione, per quanto questi possano essere suddivisi in componenti variamente specializzate che attraversano zone diverse e distanti della macchina. La condizione della ricorsività richiede che i flussi d'informazione parallela percorrano anelli chiusi di stadi mantenendo lo stato di sincronizzazione e che pertanto le connessioni tre le diverse parti del sistema siano reciprocate da connessioni che procedono in senso inverso. E' un dato di fatto che quasi tutte le connessioni tra le parti operative del cervello siano reciprocate da connessioni inverse (Asanuma e Crick, 1986) e che nel cervello esistano sequenze di stadi chiuse ad anello che potrebbero essere deputate al processamento ricorsivo dell'informazione: il circuito di Papez e quello extrapiramidale (Fig.12) sono i due più evidenti (Nieuwenhuys *et al*., 1978).

La costruzione di calcolatori massicciamente paralleli presenta tuttavia una difficoltà apparentemente insormontabile che il cervello animale sembra invece avere efficacemente risolto: un grandissimo numero di unità di calcolo deve concorrere simultaneamente allo svolgimento di uno stesso processo complessivo, la cui efficienza dipende fortemente dall'architettura del sistema. Ora, dato che l'indirizzamento dei dati è fornito dalle connessioni interne, e che l'esecuzione di un programma richiede che i dati siano indirizzati in modi imprevedibilmente vari e diversi, si comprende facilmente che l'esecuzione di un programma richiede continue modificazioni di architettura del sistema, oppure l'utilizzo, alla bisogna, di moduli variamente specializzati verso i quali i flussi d'informazione devono essere rapidamente direzionati, attivati e disattivati, dato che in generale l'architettura ottimale per un certo processo non lo è per un altro.

Nel primo caso si pone il problema di quali tecniche di controllo dei flussi d'informazione e quali metodi di rapida riorganizzazione dell'architettura possono essere usati per ottenere una grande efficienza di calcolo. Nel secondo caso, di quanto vasto debba essere il repertorio dei moduli specializzati per soddisfare alle esigenze di calcolo del sistema. Ma in entrambi i casi, possiamo dire che i programmi tendono a coincidere con l'architettura del sistema.

Se inoltre richiediamo che il calcolatore sia universale, la capacità di riconfigurare rapidamente e incessantemente la sua architettura diventa indispensabile. Un calcolatore parallelo universale non può assumere un'architettura ottimale solo per un particolare tipo di processi. Esso deve essere abbastanza versatile da eseguire una gran varietà di processi disposti in sequenze imprevedibili, e perciò di modificare in breve tempo e in modo coordinato l'architettura di tutte le sue parti. Come si può implementare l'insieme di tutte queste funzioni in una macchina parallela?

Questo problema è stato certamente risolto dai cervelli animali, dato che la selezione naturale ha provveduto da sempre alla scoperta delle soluzioni ottimali per tutte le funzioni dei sistemi biologici. Il cervello umano è certamente una macchina parallela universale capace di eseguire in pochi istanti processi ricorsivi di complessità indescrivibile. Come si potrebbe altrimenti spiegare, ad esempio, la complessità dei sogni? Come riesce dunque il cervello a fare ciò che appare tanto difficile per una macchina artificiale? Nelle parti del saggio che seguiranno cercherò di rispondere a questa domanda.

Bibliografia

1. Aiello, M., Albano, A., Attardi, G. e Montanari, U. (1976) *Teoria della computabilità, logica, teoria dei linguaggi formali*. Materiali Didattici ETS, ETS Pisa.
2. Arbib, M. (1968) *La mente, le macchine e la matematica*., Ed. Boringhieri, Bologna.

3. Asanuma, A. e Crick, F. (1986) in *Parallel Distributed Processing*, McClelland e Rumelhart Eds., Vol. 2, 333-371.
4. Barlow, H.B. (1968) in *Cybernetics*, Evans e Robertson Eds., p. 183-207, Butterworths, London.
5. Bartlett, S.J. (1992*) Reflexivity. A Source-Book in Self-Reference*. North-Holland.
6. Boole, G. (1976) *Indagine sulle leggi del pensiero*. Einaudi Ed. Torino.
7. Codenotti, B. e Leoncini, M. (1990) *Fondamenti di calcolo parallelo*. Addison-Wesley.
8. Damasio, A. R. (1994) *L'errore di Cartesio*. Adelphi Edizioni, Milano.
9. Damasio, A. R. (2000) *Emozione e Coscienza*. Adelphi Edizioni, Milano.
10. Damasio, A. R. (2004) *Alla ricerca di Spinoza*. *Emozioni, sentimenti e cervello*. Adelphi Edizioni, Milano.
11. De Luca, A. e Ricciardi, L.M. (1981) *Introduzione alla cibernetica*. F.Angeli Ed., Milano. Mazurkievicz, A. (1977), voce *Algoritmo* in *Enciclopedia*, Einaudi Ed., Torino.
12. Freeman, W.J. (1999) *Come pensa il cervello*. Einaudi.
13. Hebb, D.O. (1949) *The Organization of Behavior*. John Wiley Ed.; traduzione italiana (1975) *L'organizzazione del comportamento*, Ed. F.Angeli.
14. Hofstadter, D.R. (1985) *Gödel, Escher, Bach: un'Eterna Ghirlanda Brillante*. Adelphi Ed. Milano.
15. Hofstadter, D.R. (2008) *Anelli nell'Io. Che cosa c'è al cuore della coscienza*. Mondadori.
16. Konolige. K. (1985) *A computational Theory of Belief Introspection*. Proceedings IJCAL-85. Anche in *Reflexivity*, Bartlett Ed. (1992), pp: 343-362.
17. Libet, B., Wright, E.W.Jr., Feinstein, B. and Pearl, D.K. (1979) Subjective referral of the timing for a conscious sensory experience. *Brain*. 102:193-224.
18. McCulloch, S.W. & Pitts, W. (1943) *Bull. Math. Biophysics* , 5:115-133.
19. Nobili, R. (2009) New perspectives in brain information processing. *J Biol. Phys*. 35:347-360.
20. Von Neumann, J. (1966) *Theory of Self-Reproducing Automata*. University of Illinois Press, Urbana.
21. Nieuwenhuys, R., Voogd, J. e van Huijzen, Chr. (1978) *Sistema nervoso centrale. Testo-Atlante.* Piccin Editore, Padova.
22. Nobili, R. e Pesavento, U. (1996) Generalized von Neumann Automata. In *Artificial Worlds and Urban Studies*, Besussi e Cecchini Ed.s. Convegni DAEST, Dipartimento di Analisi Economica e Sociale del Territorio, Venezia. Nobili, R. (2004), il programma degli automi di von Neumann (WJVN.EXE) è disponibile all'indirizzo http://pd.infn.it/~rnobili
23. Putnam, H. (1979) voce *Logica* in *Enciclopedia*, Einaudi Ed., Torino.
24. Singer, W. (1999) Neural synchrony: a versatile code for the definition of relations? *Neuron*, 24:69-75.
25. Shannon, C. (1971) *La teoria matematica delle comunicazioni*. Ed. Eta Compas.
26. Tarski, A. (1966) Verità e dimostrazione. In *Verità e dimostrazione*. *Questioni di matematica*. A cura di C.Ciliberto, Lettura da Le Scienze, (1978) Le Scienze S.p.A editore, Milano.
27. Turing, A.M. (1936), On Computable Numbers, with an Application to the Entscheidungsproblem. *Proceedings of the London Mathematical Society*, 2 42: 230–265; (1937) On Computable Numbers etc.: A correction", *Proceedings of the London Mathematical Society*, 2 43: 544–46.

Padova, 17 Marzo 2010.